Kartta: Extracting Landmarks near Personalized Points-of-Interest from User Generated Content

Arttu Perttula
Tampere University of Technology
PO Box 300, FI-28101, Pori, Finland
arttu.perttula@tut.fi

Scott Carter, Laurent Denoue
FX Palo Alto Laboratory, Inc.
3400 Hillview Ave., Bldg. 4, Palo Alto, CA 94304
carter,denoue@fxpal.com

ABSTRACT

Most mobile navigation systems focus on answering the question, "I know where I want to go, now can you show me exactly how to get there?" While this approach works well for many tasks, it is not as useful for unconstrained situations in which user goals and spatial landscapes are more fluid, such as festivals or conferences. In this paper we describe the design and iteration of the Kartta system, which we developed to answer a slightly different question: "What are the most interesting areas here and how do I find them?"

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: Miscellaneous

Keywords

Mobile, navigation, media

1. INTRODUCTION

Navigating an unfamiliar, eventful environment such as a festival or conference can be challenging. It can be hard to decide what to see, and when you have decided where to go, it can be difficult to know how to get there. While several research projects have investigated the mobile navigation of physical environments, past work has focused on extracting general landmarks from relatively static sources, such as structured maps. In this paper we describe a system, Kartta, that derives points-of-interest and associated landmarks from user generated content captured onsite. With this approach, Kartta helps users navigate standard environments as well as temporary events, such as festivals and fairs. Using social network and preference information, the system also reflects a more personalized set of points-of-interest. Furthermore, unlike past work, the system supports a diversity of landmark representations, including text tags, images, audio and video clips.

2. KARTTA

Kartta is composed of a server and a mobile application built using Mobile Python. The mobile application can be used to capture content and context data which are sent to the server automatically. The server uses this information to create a map of the immediate region around a user highlighting points-of-interest as well as landmarks to help the user navigate. The mobile application polls the server regularly for a new map and set of landmarks.

The mobile application visualizes not only the augmented map returned by the server but also the users' current location (Figure 1). Users can configure the interface either to show media captured by an explicit list of friends, or to show only their own captures. Because, as Fleck et al. point out [2], mobile users often feel that capturing media is too distracting, the application includes a simple, one-button interface for voting a current location up or down. A positive vote is represented on the map as a green dot and a negative vote a red dot. Over time, votes will create implicitly an interest-map of a place.

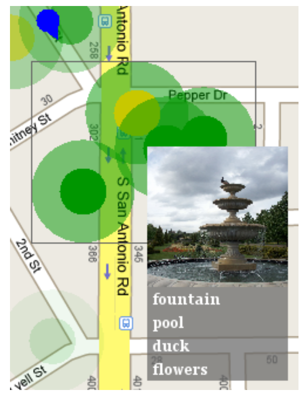

Figure 1: The Kartta interface showing a user (blue dot, upper left) making her way toward a cluster of interest areas. Kartta chose the image shown automatically as the best media to direct users toward this cluster (it is associated with the interest area with a yellow center). Text below the image are unique tags from interest areas in the cluster.

Mobile users can also launch media capture applications with another button. The mobile application sends media to the server along with context information (if available) including phone tilt (up, level, or down), compass orientation, time-of-day, and location. Context data can also be recorded continuously. The application currently senses location information via either embedded or attached GPS devices. While GPS currently affords only a gross estimate of location, we believe that it is sufficient since our application depends on aggregated data (votes). Furthermore, users can add tags to disambiguate areas of interest. Finally, media captured at a location automatically correspond to a positive vote for that place (similar to the approach in [4]).

The server ingests and clusters media and votes sent from mobile devices, selects representative captures for each cluster, and generates routes (see Figure 2). The server saves mobile media in a database, and a set of ingest and access services produce multiple downsampled representations of

Copyright is held by the author/owner(s).
MobileHCI'09 September 15-18, 2009, Bonn, Germany
ACM 978-1-60558-281-8.

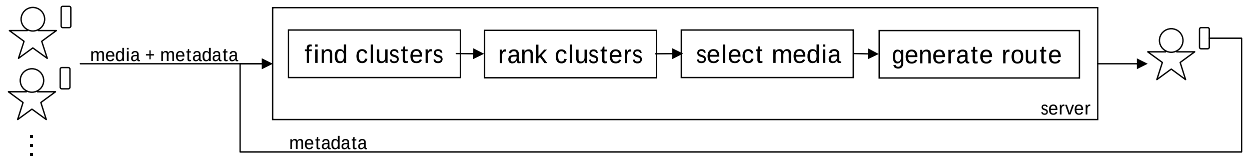

Figure 2: Data flow in the Kartta system. Using the mobile application, participants upload media, votes, and associated metadata to the server. Synchronously or asynchronously, other participants request a map from the server, optionally sending location and orientation information. The server finds and ranks clusters, selects media within each cluster, generates routes, and finally returns this information to the mobile client.

uploaded media: thumbnails of photos, keyframes and lower resolution copies of videos, and copies with lower sample rates for audio clips. The mobile application can use these different representations to provide tiered access to media.

To organize captured media, the server first finds and ranks clusters of interest areas, ranks media within each cluster, and then generates route information.

3. EVALUATION

We first built a simulation tool to tune weights in our ranking algorithm. After running a series of simulations, we ran a pilot field experiment following Massimi et al.'s scavenger hunt approach for evaluating mobile collaborative systems [5]. Also, inspired by Baus' et al.'s finding that audio can be useful for navigation tasks [1], we included a variety of media types in our experiment. We recruited six single users from our lab to use Kartta to find and record objects of interest in a semi-familiar, semi-urban environment. We pre-recorded objects using a mixture of media. We then gave participants a device with the Kartta application and asked them to take a photo of the object they thought was being recorded or tagged at each location (for example, one hotspot was linked to an audio clip of the chimes of a clock). Media captured by participants was saved to the device locally and uploaded to the server in a separate thread. Participants were also asked to avoid areas marked with negative votes. One researcher shadowed the participant both to record behavior and to answer questions about the interface in case the participant got stuck. At the end of the task, we asked participants several follow-up questions designed to determine which media was most helpful for navigation, the ease of navigating the map, and their ability to avoid negative areas.

3.1 Results

Despite there being no incentive, all participants completed the study in the time allotted. However, only one participant successfully captured photos of all objects, and one captured only five. On average, participants captured eight out of ten objects. Media in the set that only appeared after the start of the experiment were those most likely to be left out. All participants chose photos as either the first or second most useful media for navigation. Four chose tags as the first or second most useful. One participant commented that it was easier to find those media that he could at least "roughly make out from a distance." Overall, participants found it useful that the interface reflected their orientation as well as location (four on a five-point scale on average), and they found it easy to navigate the map (also four out of five). In general, participants had fun with the task, one noting that it was "like virtual geocaching."

3.2 Discussion

Our overall experience with the simulation and field experiment indicate that user generated content can help people navigate unfamiliar events. In particular, participants seemed to find it natural to move toward "hot spots" on the map given rich representations of perceptible landmarks. However, more work is necessary to support unconstrained environments. For example, some sensors may be available only sporadically. In these cases, the server should automatically adjust the weights used to calculate cluster importance. We were also surprised that users in the field study preferred photos and tags considerably more than audio and video.

4. RELATED WORK

Some systems support subsets of the Kartta platform. Jaffe et al. built a system that uses contextual information to select key photos from a collection [4]. However, their work focuses on summarization rather than navigation. Grabler et al. developed a prototype that automatically generates maps using building textures, road geometry, and external landmark information [3]. In this work, the authors focus on static structures and landmarks rather than temporary environments. Finally, studies by Baus et al. showed that audio landmarks can aid navigation tasks, but their work did not involve user generated content [1].

5. CONCLUSION AND FUTURE WORK

While our experimentation is limited, we have shown that it is possible to gather interest areas and landmarks from user generated data. Also, because of our focus on fluid, informal environments we made design decisions that allow users to derive some value from the system with only minimal input, including one-click location voting, automatically capturing metadata, and heatmap visualizations. After running more complex simulations, the next step for this work is a full-fledged field experiment with a group or friends or colleagues in an unconstrained environment.

6. REFERENCES

[1] J. Baus et al. Auditory perceptible landmarks in mobile navigation. In *IUI '07*, pp. 302–304.

[2] M. Fleck et al. From informing to remembering: Ubiquitous systems in interactive museums. *IEEE Pervasive Computing*, 1(2):13–21, 2002.

[3] F. Grabler et al. Automatic generation of tourist maps. In *Siggraph '08*, pp. 1–11.

[4] A. Jaffe et al. Generating summaries and visualization for large collections of geo-referenced photographs. In *MIR '06*, pp. 89–98.

[5] M. Massimi et al. Scavenger hunt: An empirical method for mobile collaborative problem-solving. *IEEE Pervasive Computing*, 6(1):81–87, 2007.